Efficiency, selectivity and robustness of nucleocytoplasmic transport

A. Zilman [1], S. Di Talia [1], B. T. Chait [3] , M. P. Rout [2], and M. O. Magnasco [1]

[1] *Laboratory of Mathematical Physics*,

[2] *Laboratory of Cellular and Structural Biology*

[3] *Laboratory of Mass Spectroscopy and Gaseous Ion Chemistry*

The Rockefeller University, 1230 York Ave, New York, NY 10021

Running title: Physical model of transport through the NPC.

Abstract

All materials enter or exit the cell nucleus through nuclear pore complexes (NPCs), efficient transport devices that combine high selectivity and throughput. A central feature of this transport is the binding of cargo-carrying soluble transport factors to flexible, unstructured proteinaceous filaments called FG-nups that line the NPC. We have modeled the dynamics of transport factors and their interaction with the flexible FG-nups as diffusion in an effective potential, using both analytical theory and computer simulations. We show that specific binding of transport factors to the FG-nups facilitates transport and provides the mechanism of selectivity. We show that the high selectivity of transport can be accounted for by competition for both binding sites and space inside the NPC, which selects for transport factors over other macromolecules that interact only non-specifically with the NPC. We also show that transport is relatively insensitive to changes in the number and distribution of FG-nups in the NPC, due mainly to their flexibility; this accounts for recent experiments where up to half of the total mass of the NPC has been deleted, without abolishing the transport. Notably, we demonstrate that

previously established physical and structural properties of the NPC can account for observed features of nucleocytoplasmic transport. Finally, our results suggest strategies for creation of artificial nano-molecular sorting devices.

Introduction

The contents of the eukaryotic nucleus are separated from the cytoplasm by the nuclear envelope. Nuclear pore complexes (NPCs) are embedded in the nuclear envelope and are the sole means by which materials exchange across it. Water, ions, small macromolecules (smaller than 40 kDa) [1] and small neutral particles (of diameter less than $5-6$ nm) can diffuse unaided across the NPC [10], while larger macromolecules (and even many small macromolecules) will generally only be transported efficiently if they display a signal sequence, such as nuclear localization signals (NLSs) or nuclear export signals (NESs), that bind to cognate soluble transport factors, which facilitate the passage through the NPC. The resulting transport factor-cargo complexes then traverse the NPC. The best studied transport receptors belong to a family of structurally related proteins, collectively termed β-karyopherins, although other transport factors can also mediate nuclear transport, particularly the export of mRNAs (reviewed in [1–3, 5]). NPCs can pass cargos up to 30 nm diameter (such as mRNA particles) , at rates as high as several hundred macromolecules per second, each one dwelling in the NPC for only several milliseconds [7, 8].

In this paper, we shall focus on karyopherin-mediated import, although our conclusions pertain to other types of nucleocytoplasmic transport as well, including mRNA export. During import, karyopherins bind cargoes in the cytoplasm via their NLSs. The karyopherin-cargo complexes then translocate through NPCs to the nucleoplasm, where the cargo is released from the karyopherin by a nuclear enzyme, RanGTP. The high affinity of RanGTP binding to karyopherins allows it to displace the cargoes from the karyopherins. Subsequently, karyopherins with bound RanGTP travel back through the NPC to the cytoplasm, where conversion of RanGTP to RanGDP is catalyzed by the cytoplasmic enzyme RanGAP. The energy released from GTP hydrolysis is used to dissociate RanGDP from the karyopherins,

which are now ready for the next cycle of transport. Importantly, this GTP hydrolysis is the only step in the process of nuclear import that requires input of metabolic energy. Overall, the energy obtained from RanGTP hydrolysis is used to create a concentration gradient of karyopherin-cargo complexes between the cytoplasm and the nucleus, and the process of actual translocation across the NPC occurs purely by diffusion (reviewed in [1–6]).

Conceptually, nuclear import can be divided into three stages: first, the loading of cargo onto karyopherins in the cytoplasm; then, the translocation of karyopherin-cargo complexes through the NPC (cf. Fig. 1), and finally the release of cargo inside the nucleus. The first and last stages have been the subject of numerous studies, and are relatively well understood, being soluble-phase reactions amenable to biochemical characterization (reviewed in [1–5]). The intermediate stage of transport is much less understood. Nevertheless, it is clear that the ability of karyopherins (and other transport factors) to bind a particular class of NPC-associated proteins, known collectively as FG-nups, is a key feature of the transport process, and allows them to selectively and efficiently pass with their cargoes through the NPC. In particular, experiments in which the FG-nup-binding sites on the karyopherins were mutated [9] show that disrupting the binding of karyopherins to FG-nups impairs transport. Current estimates of the binding affinity of karyopherins to the FG-nups are in the range 10-1000 nM, (or $5 - 10 k_B T$ per binding site) depending on the FG-nup and karyopherin type [12, 13, 15].

Transport through the NPC recently gave rise to several theoretical hypotheses [6, 14, 29–31]. Here we develop a theory to explain the mechanism of the intermediate stage of nucleocytoplasmic transport - translocation through the NPC - which provides a physical basis for our virtual gating model [2, 6].

A theory of NPC-mediated transport has to answer several major questions: *i*) The transport of any macromolecules, even those potentially able to negotiate the NPC unaided, is increased by their specific association with NPCs [11]. How does the NPC achieve high transport efficiency of the cargoes of variable sizes and in both directions, through only passive diffusion of the transport factor-cargo complexes? *ii)* How does binding of transport factors to FG-nups facilitate transport efficiency while maintaining a high throughput (up to

hundreds of molecules per second per NPC) [2, 7, 8]? *iii)* NPCs largely exclude non-specific macromolecules in favor of transport factor-bound cargoes (e.g., [1]). How is this high degree of selectivity achieved? *iv)* Neither deletion of up to half of the mass of the FG-nups, nor deletion of asymmetrically disposed FG-nups that potentially set up an affinity gradient, abolish transport [17]. Directionality of transport across the NPC can even be reversed by reversing the concentration difference of the RanGTP [18]. How can we account for such a high degree of robustness?

The model we present here provides answers to these questions, explaining the functional features of the NPC in terms of its known structural and physical properties. We show how karyopherin binding to the flexible filaments inside the NPC gives rise to efficient transport. We show that the competition for the limited space and binding sites within the NPC leads to highly selective filtering. Finally, we show how the flexibility of the FG-nups accounts for the high robustness of NPC-mediated transport with respect to structural changes [17]. We conclude by summarizing the salient features of nuclear transport, and discuss verifiable experimental predictions of the model.

Setting up a physical model of NPC transport

The NPC is a protein assembly spanning the nuclear envelope that contains a central channel, about 35 nm in diameter, connecting the nucleoplasm with the cytoplasm. The internal volume of this channel, as well as large fractions of the nuclear and cytoplasmic surfaces of the NPC, are occupied by unstructured polypeptide chains collectively known as FG-nups because they carry unusually high numbers of repeats containing phenylalanine-glycine pairs. FG-nups are known to be flexible [19–21]. Since the FG-nups also protrude into the nucleus and the cytoplasm, the effective length of the nuclear pore complex is estimated to be 70 nm [1–3]. The details of the distribution of the FG-nups inside the central channel and the external surfaces of the NPC, as well as the exact number of binding sites on the FG-nups and on the karyopherins, are not fully known. We shall therefore make no specific assumptions about the distribution of FG-nups, interactions between them, and their density, degree of

flexibility or conformation within the NPC.

We represent the transport through the NPC as a combination of two independent processes contributing to the movement of the karyopherin-cargo complexes through the central channel of the NPC: *(i)* the binding and unbinding of the karyopherins to the FG-nups, and textit(ii) the spatial diffusion of the complexes, either in the unbound state, or while still bound to a flexible FG-nup [3]. The complexes entering the NPC from the cytoplasm stochastically hop back and forth inside the channel until they either reach the nuclear side, where the cargo is released by RanGTP, or return to the cytoplasm. Detachment from the FG-nups and exit from the NPC can be either thermally activated, or catalyzed by RanGTP directly at the nuclear exit of the NPC [1]. A schematic illustration of transport through the NPC is shown in Fig. 1.

Transport efficiency arises from the karyopherins' ability to bind to FG-nups.

It is important to distinguish between two different properties of the transport process: the *speed* and the *probability* with which individual complexes that enter from the cytoplasm traverse the NPC to nuclear side. [22–24]. As we show below, binding of karyopherins to the FG-nups increases their transport efficiency, i.e., the probability of them traversing the NPC; in the absence of such binding, the probability of traversing the NPC is low.

For simplicity, we shall assume that the unbinding and rebinding occur faster than lateral diffusion of karyopherin-cargo complexes along the channel(our conclusions nevertheless were verified by computer simulations for any ratio of binding-unbinding rate to diffusion rate). In this limit, the hopping between discrete binding sites can be approximated by diffusion in an effective potential which combines two effects: entropic repulsion of the FG-nups, as the large cargoes have to compress and displace them to enter the channel, and attraction due to binding to the binding sites in the FG-nups [2, 6, 21]; this approximation is justified below. Thus, we represent the transport of karyopherin-cargo complexes through the NPC as a one-dimensional diffusion in a potential $U(x)$ (expressed in units of k_BT), in the interval

$0 < x < L$, as illustrated in Fig. 2. We solve the model for an arbitrary shape of the potential, determined by the distribution of the binding sites on the FG-nups along the channel. We do not directly model the diffusion of complexes outside of the NPC. Instead, we assume that karyopherin-cargo complexes stochastically enter the NPC from the cytoplasm, with an average rate J at $x = R$; J is proportional to the concentration of the karyopherin-cargo complexes in the cytoplasm [16]. The release of karyopherin-cargo complexes from FG-nups by RanGTP at the nuclear exit is modeled by imposing an exit flux J_e at $x = L - R$, as shown in Fig. 2. Thus, the length of the NPC corresponds to the interval from $x = R$ to $x = L - R$; the regions of length R (of the order of the width of the channel [16]) at both ends of the interval correspond to the distance outside the NPC over which the particles diffuse away by three-dimensional diffusion into either the nucleoplasm or the cytoplasm. We use absorbing boundary conditions at $x = 0$ and $x = L$ that correspond to a karyopherin-cargo complex returning back to the cytoplasm, or going through to the nucleus, respectively.

We neglect the difference in the diffusion coefficient of the complexes inside and outside the NPC in order to focus on the role of karyopherin binding in the import process. We also assume that no current enters from the nucleus, as the cargoes are released from the karyopherins in the nucleus by RanGTP. Finally, we neglect the variations of the potential in the direction perpendicular to the channel axis. The effect of these factors will be studied in detail elsewhere. Under the above assumptions, transport of the karyopherin-cargo complexes through the NPC is described by the diffusion equation for the density of the complexes inside the channel, $\rho(x)$ [22]:

$$\frac{\partial \rho(x)}{\partial t} = -\frac{\partial J(x)}{\partial x} \tag{1}$$

where the local current $J(x)$ is given by:

$$J(x) = -De^{-U(x)}\frac{\partial}{\partial x}[e^{U(x)}\rho(x)] \tag{2}$$

where D is the diffusion coefficient.

The entrance current J splits into J_0, and J_M, corresponding to the flux of complexes

returning to the cytoplasm and going through to the nucleus, respectively. At a position $x = L-R$, the transmitted flux J_M splits into J_e and J_L, which correspond to the karyopherin-cargo complexes released from the FG-nups by RanGTP and to thermally activated release, respectively (cf. Fig. 2).

The steady state solution of eqs.(1,2) with entrance flux J and satisfying the boundary conditions $P(0) = P(L) = 0$, is

$$\begin{aligned} \rho(x) &= |J_0|\frac{1}{D}e^{-U(x)}\int_0^x e^{U(x')}dx' \quad \text{for } 0 < x < R \\ \rho(x) &= \frac{1}{D}e^{-U(x)}\left[|J_0|R - J_M\int_R^x dx' e^{U(x')}\right] \\ &\quad \text{for } R < x < L-R \\ \rho(x) &= J_L\frac{1}{D}e^{-U(x)}\int_x^L e^{U(x')}dx' \ \text{ for } L-R < x < L \end{aligned} \tag{3}$$

The sum of the flux of karyopherin-cargo complexes going through the NPC and of those returning to the cytoplasm is equal to the total flux of complexes entering the NPC; hence $|J_0| + J_M = J$; similarly, $J_M - J_L = J_e$. The flux J_e is proportional to the number of complexes present at the nuclear exit, and to the frequency $J_{\rm ran}$ with which RanGTP molecules hit the nuclear exit of the NPC: $J_e = J_{\rm ran}\rho(x = L-R)R$. Recalling that that the potential outside the channel is zero ($U(x) = 0$) for $0 < x < R$ and $L-R < x < L$, and using the continuity of $\rho(x)$ at $x = L-R$, one obtains for P_{tr}, the probability of a given karyopherin-cargo complex reaching the nucleus:

$$P_{tr} = J_M/J = \frac{1}{2 - K/(1+K) + \frac{1}{R}\int_R^{L-R} dx e^{U(x)}} \tag{4}$$

where $K = (J_{ran}R^2/D)\, e^{-U(L-R)}$.

Eq.(4) is the main result of this section and has several important consequences. The probability P_{tr} defines the transport efficiency, which is seen to increase with the potential

depth E, (defined as $E = -\min_x U(x)$, Fig. 2), proportional to the binding strength of the karyopherins to the FG-nups. In the absence of binding, P_{tr} is small ($\sim R/L$), so that a complex will on average return to the cytoplasm soon after entering the NPC. An attractive potential inside the NPC *increases* the time the complex spends inside the NPC and thus *increases* the probability that it reaches the nuclear side, rather than returns to the cytoplasm.

When RanGTP only releases cargo from its karyopherin, but not from the FG-nups (i.e. $J_{\rm e} = 0$), the maximal translocation probability P_{tr} is 1/2. However, in the case when RanGTP also releases karyopherin-cargo complexes from FG-nups, the translocation probability P_{tr} can reach 1. Importantly, the latter effect is more pronounced for large K, that is for strong binding at the exit.

The second important consequence of eq. (4) is that the translocation probability P_{tr} depends only weakly on the *shape* of the potential $U(x)$. Thus, we predict that the transport properties of the NPC are relatively insensitive to details of the distribution of FG-nups inside the NPC, and to the distribution of the binding sites on FG-nups.

Limitations of space and number of binding sites provide a mechanism for selectivity

The discussion of the previous section neglects the mutual interactions between karyopherin-cargo complexes. However, a large interaction strength E increases transport efficiency at the expense of an increased transport time $T(E)$ (which grows roughly exponentially with E [22]), leading to accumulation of particles inside the channel. Therefore, for high E, interactions between the complexes due to steric occlusion and decrease in the available binding sites become important.

To quantitatively investigate how mutual interference between translocating karyopherin-cargo complexes affects transport efficiency, we performed dynamic Monte Carlo simulations of the diffusion of complexes inside the NPC, in the potential $U(x)$, using a variant of the Gillespie algorithm [25–27]. The simulations are a discrete version of the continuum formulation of the previous section. The interval $[0, L]$ is represented by N discrete "sites"; these sites do

not represent the actual binding sites, but correspond to the length of a diffusion step. For simplicity, we allow only one particle at each site at any moment of time, which models mutual occlusion between complexes; however, our results pertain for any allowed occupancy. In line with our analytical model above, karyopherin-cargo complexes are deposited at the site i_R, if it is unoccupied, with a probability $J\frac{L^2}{DN^2}$ per simulation step. When a complex reaches site $i = 0$ (cytoplasm) or $i = N$ (nucleus), it is removed. The complexes present at site $i = N - i_R$ can be removed directly, with the probability $J_{\rm ran}\frac{L^2}{DN^2}$, which models the effect of the release of the complexes from FG-nups by nuclear RanGTP. Once inside the channel, a complex present at site i can hop to an adjacent unoccupied site $i \pm 1$ with the following probability:

$$P(i \to i \pm 1) = r_{i,i\pm1}/(J + J_{ran}x_{N-i_R} + \sum_{i=1}^{N} r_{i,i\pm1}) \tag{5}$$

where x_i is the site occupancy: $x_i = 0$ if the site is unoccupied, and $x_i = 1$ if a particle is present at the site; $r_{i,i\pm1} = \frac{D}{(L/N)^2}\exp((U_i - U_{i\pm1})/2)x_i(1 - x_{i\pm1})$ is the transition rate from site i to $i \pm 1$ [25–27].

The results of our simulations are shown in Fig. 3. For low interaction strength E the translocation probability curves for all entrance fluxes J collapse onto a single line which is predicted by the analytical solution of the previous section, because in this regime there are few complexes simultaneously present in the channel, and the interactions between them are negligible. For stronger binding, karyopherin-cargo complexes accumulate inside the channel, blocking the inflow of additional complexes.

The main conclusion of the simulations, as shown in Fig.3 **A**, is that the transport through the NPC is maximal for an optimal value of the interaction strength E_c, which depends on the entrance rate J. The decrease in NPC throughput at high E becomes significant when the number of karyopherin-cargo complexes in the channel reaches a certain critical value, M_c. This critical occupancy of the channel M_c is proportional both to the entrance flux J, and to the residence time $T(E_c)$, $M_c = JT(E_c)$, and therefore the optimal interaction strength is higher for low entrance fluxes as illustrated in panel **B** of Fig. 3 .

Existence of an optimal interaction strength provides a mechanism for the selectivity of

NPC-mediated transport. The translocation probability is high for a particular strength of interaction of karyopherins with FG-nups, while macromolecules that do not interact with FG-nups are filtered out.

Competition between non-specifically binding macromolecules and karyopherins enhances the selectivity of the NPC

As shown in the previous section, the binding of karyopherins to FG-nups provides a mechanism of selectivity, because transport of the karyopherin-cargo complexes is highest for a specific value of their interaction strength with the FG-nups. However, the maximum depicted in Fig. 3**A** is broad; the translocation probability is significant even for binding strengths considerably lower than the optimal one. For instance, if the optimal interaction strength is $E = 10kT$, macromolecules whose interaction strength is $5kT$ a have probability of reaching the nucleus that is just a half of the optimal one. This broad maximum allows NPC-mediated import to function efficiently across a broad range of transport factor binding strengths. On the other hand, this might also permit passage of macromolecules that bind non-specifically to FG-nups (e.g., due to electrostatic interactions). However, proper functioning of living cells requires a high selectivity of the NPC - how might this be achieved?

So far, Fig. 3 only takes into account the competition between complexes of identical binding strength for space inside the channel. However, in a situation where optimally-binding karyopherins compete for space and binding sites inside the channel with weakly-binding macromolecules, the passage of the latter is sharply reduced which sharply increases the selectivity of the NPC. Qualitatively, because the residence time is higher for strongly-binding karyopherins, a weakly-binding macromolecule upon entering the channel will with high probability find it blocked by a karyopherin. Thus, the low affinity macromolecule will with high probability return to the cytoplasm due to a relatively low residence time. On the other hand, if a karyopherin enters a channel that is already occupied by another karyopherin, there is a high probability that it will reside inside the channel long enough for the first karyopherin to get through. As free karyopherins exchange back and forth across the NPC constantly, there

will always be karyopherins (or other transport factors) binding in the NPC and so excluding non-specific macromolecules, making the NPC a remarkably efficient filter.

These heuristic arguments were verified via computer simulation. Two species of particles of different binding affinities (representing the karyopherins and another macromolecule that can non-specifically bind to the FG-nups), are deposited stochastically at the NPC entrance with the same average rate J. As in the previous section, the particles diffuse inside the channel until they either reach the nucleus, or return to the cytoplasm; a site can be occupied by only one particle.

The results of the simulations are presented in Fig. 4, which compares the transport efficiency of strongly binding macromolecules to that of weakly binding ones in the case when they compete for the space inside the NPC. The competition for the space inside the channel between the translocating particles dramatically narrows the selectivity curve as compared with Fig. 3. This effect is a novel mechanism for the enhancement of transport selectivity beyond what is expected from the binding affinity differences alone. In contrast to other mechanisms of specificity (e.g. kinetic proofreading [28]), no additional metabolic energy is required for this enhanced discrimination. Instead, selectivity is achieved by competition producing a differential NPC response to two ranges of binding affinities. In the range of higher binding affinities occupied by transport factors, passage across the NPC is efficient; whereas in the low range of affinities, transmission is effectively prevented.

The flexibility of FG-nups accounts for the robustness of NPC-mediated transport

In the previous sections, we have used a continuous potential to model transport through the NPC. However, in reality the translocating karyopherin-cargo complexes hop between discrete binding sites that are located on the flexible FG-nups, which fluctuate in space around their anchor point due to thermal motion [19, 21]. This allows complexes to diffuse along the channel while remaining bound to an FG-nup. A complex can also unbind from an FG-nup and rebind again to the same or a neighboring FG-nup, moving while unbound by passive diffusion. In

this section, we elucidate how the *number* of FG-nups inside the channel affects transport.

The translocation of the karyopherin-cargo complexes through the NPCcan be described as a diffusion in an array of potentials, as illustrated in Fig. 5, where each potential well $U_i(x)$ represents an FG-nup. The shape of each well depends on the binding strength of the karyopherin-cargo complex and the rigidity and the length of an FG-nup, which determine the cost of its entropic stretching in the process of spatial diffusion [21]. The blue line corresponds to the unbound state. Although for the purposes of illustration all the wells are shown to have the same form, the subsequent results are valid for arbitrary distribution of potential shapes. We shall denote the density of the karyopherin-cargo complexes in the i-th well as $\rho_i(x)$ and the density of unbound complexes as $\rho_0(x)$. The lateral diffusion of the complexes, combined with the binding and unbinding to the FG-nups is then described by the following equations [34]:

$$\begin{aligned}
\frac{\partial \rho_0(x)}{\partial t} &= \frac{\partial}{\partial x} e^{-U_0(x)} \frac{\partial}{\partial x} e^{U_0(x)} \rho_0(x) + \\
&+ \sum_i [r_{i0}(x)\rho_i - r_{0i}(x)\rho_0(x)] \\
\frac{\partial \rho_i(x)}{\partial t} &= \frac{\partial}{\partial x} e^{-U_i(x)} \frac{\partial}{\partial x} e^{U_i(x)} \rho_i(x) + \\
&+ r_{0i}(x)\rho_{0i} - r_{i0}(x)\rho_i(x)
\end{aligned} \tag{6}$$

where $r_{i0}(x)$ and $r_{0i}(x)$ are the local unbinding and binding rates, respectively. They are related by the detailed balance condition: $r_{0i}(x)/r_{i0}(x) = e^{-U_i(x)+U_0(x)}$. If the unbinding rates are fast compared to the diffusion time across the wells (i.e. $r_{i0}^{-1} e^{U_0 - U_i} \gg D/(2p^2)$), the relative densities of bound and unbound complexes are at their local equilibrium Boltzmann ratio [34]: $\rho_i(x) = \frac{e^{-U_i(x)}}{\sum_i e^{-U_i(x)}} \rho(x)$ where $\rho(x) = \sum_i \rho_i(x)$ is the total density of the complexes at a position x. Summing up the equations (6), one obtains an equation for the total complex density $\rho(x)$:

$$\frac{\partial \rho(x)}{\partial t} = \frac{\partial}{\partial x} e^{-U_{\rm eff}(x)} \frac{\partial}{\partial x} e^{U_{\rm eff}(x)} \rho(x) \tag{7}$$

where $U_{\rm eff} = -\ln(e^{-U_0(x)} + \sum_i e^{-U_i(x)})$ [34]. Thus, the process of translocation through the NPC can be described as simple lateral diffusion in the effective potential $U_{\rm eff}$, which leads to the eqs. (1,2).

As we have shown in previous sections, the transport properties of the NPC are relatively insensitive to the detailed shape of the effective potential. Thus, we predict that the transport trough the nuclear pore is robust with respect to the variations in the number of the FG-nups, as indeed observed experimentally [17].

Even more strikingly, the transport properties of the NPC are not sensitive to the number of the FG-nups also in the case when the FG-nups are distributed sparsely inside the NPC, without a significant overlap. We illustrate this issue through a marginal case where the fluctuation regions of each FG-nup barely touch, represented by the potential shown by the black line in Fig. 6**A**. The flat central part of each potential well corresponds to the diffusion of the karyopherin-cargo complex while bound to an FG-nup, and the sharply rising regions at the borders correspond to unbinding of the complex and its transfer to the next filament. Narrow wells correspond to relatively rigid filaments, while wide wells correspond to flexible filaments that can stretch to a long distance without significant entropic cost. The potential wells can have different widths p_i so that $\sum_{i=1}^{n} p_i = L - 2R$. All the potential wells have the same shape, U_0, re-scaled to the width of an individual well, so that the potential at a point $x = \sum_{j<i} p_j + \Delta x$ is: $U(x) = U_0(\Delta x/p_i)$.

Crucially, the transport properties of the potential shown in Fig. 6**A** *do not depend on the number of wells*. Both the translocation probability and the residence time are equivalent for the multi-well potential shown in black, and the single well potential shown in red, obtained by re-scaling an individual black well to the whole length $L - 2R$ of the channel. Indeed, it follows from eq. (4) that the translocation probability for the multi-well potential with n

wells, is

$$\begin{aligned} P_{tr} &= \frac{1}{2 - K/(1+K) + \frac{1}{R}\sum_{i=1}^{i=n-1}\int_0^{p_i} e^{U_0(x/p_i)}dx} \\ &= \frac{1}{2 - K/(1+K) + \frac{\sum_i p_i}{R}\int_0^1 e^{U_0(y)}dy} \end{aligned} \tag{8}$$

which is independent of the number and the width of the wells because $\sum_i p_i = L - 2R$. We prove in the Supporting Information that the residence time is similarly independent of the number of wells. Since both translocation probability and residence time are independent of the number of wells, the transport properties do not depend on the number of wells even for high entrance rate or binding strength, when jamming is important, as verified by computer simulation (Fig. 6**B**).

This result highlights the robustness of our model of NPC transport; in multiwell potentials of this type, the NPC's transport properties do not depend on the specific number of FG-nups, so long as they are flexible enough for their fluctuation regions to overlap, permitting complexes to freely transfer from one filament to the next.

Discussion

We have presented a physical model of transport through the NPC. The physics of diffusion in a channel and binding to flexible filaments accounts for the major observed features of NPC transport, without the need for further detailed assumptions such as the conformation and distribution of FG-nups inside the channel or the shape of the channel itself.

A crucial component of our model is the interaction between karyopherins and FG-nups, which provides a mechanism for both efficiency and selectivity. We show that in the absence of binding, the probability of crossing the NPC is much lower than the probability of returning to the cytoplasm. The binding of karyopherins to FG-nups increases the residence time of the karyopherin-cargo complexes within the NPC, preventing their premature return to the cytoplasm, and hence increasing their chance of diffusing to the nucleus, i.e., increasing the efficiency of their transport.

We show that the selectivity of the NPC arises from a balance between the *probability* and the *speed* of transport of individual karyopherin-cargo complexes. Similar ideas have been suggested to account for the transport properties of ion channels and porins [22–24, 32]. In our model, the probability of a karyopherin-cargo complex reaching the nucleus increases with the binding strength to FG-nups, but at the expense of increasing residence time inside the NPC; eventually, complexes spend so much time in the NPC that they impede the passage of other complexes through channel. Therefore, there is an optimal value of the binding strength of karyopherins to FG-nups that maximizes their transport efficiency through the NPC. This optimal binding strength is higher for low entrance fluxes J, because at low fluxes the accumulating complexes can reside longer in the channel without blocking it (Fig. 3). This observation could account for the proliferation of different karyopherins types; the binding strength of each karyopherin type might be related to the total relative abundance of its cargoes. Our model applies to both export and import processes, and explains how highly efficient nucleocytoplasmic transport can be achieved in both directions by pure diffusion.

For proper cell functioning, the NPC should selectively transport karyopherin-cargo complexes, effectively filtering out any macromolecules that do not bind specifically to the FG-nups. However, the basis for discrimination between specifically and non-specifically binding macromolecules may be as little as a few k_BT. We propose a novel mechanism that further enhances the specificity of NPC transport. This mechanism relies on the direct competition between karyopherins and non-specifically binding macromolecules; they compete for space and binding sites in the channel. As a consequence of their stronger binding, karyopherins have a longer residence time within the channel as compared with non-specifically binding macromolecules which are therefore out-competed for space and binding sites within the channel. The constant flux of cargo bound- or free karyopherins between the nucleus and cytoplasm therefore effectively excludes non-specifically binding macromolecules from the channel. Remarkably, although no metabolic energy is expended in this filtering process [28], the resulting selectivity is much higher that might be expected from only the binding affinities differences between specific and non-specific macromolecules (Fig. 4).

We have also shown that transport efficiency is enhanced when RanGTP directly releases karyopherins from their binding sites on FG-nups [1, 12] at the NPC exit - an enhancement that increases with the binding strength at the nuclear exit. Physically, high affinity binding sites at the nuclear exit of the NPC decrease the probability of return, once a complex has reached the nuclear side. This prediction may account for the observed high affinity binding sites that are localized at the nuclear side on the NPC in import pathways and cytoplasmic side in export pathways (reviewed in [2]).

Although the transport properties of the NPC depend strongly on the magnitude of the interaction strength, we predict that transport depends only weakly on spatial variations of the binding sites along the channel. Most importantly, the *number* of flexible FG-nups inside the NPC does not significantly affect transport, as long as their fluctuation regions can overlap (Figs. 5 and 6). This prediction could account for recent experiments in which up to half the total mass of FG-nup flexible regions in NPCs were deleted without seriously hampering nucleocytoplasmic transport [17]. Moreover, we predict that a gradient of binding affinity across the NPC does not, by itself, increase the throughput compared to a uniform distribution of the same sites, explaining how transport can even be reversed across the NPCby reversing the gradient of RanGTP [18]. Only in combination with a high affinity trap at the exit, and high Ran activity in releasing the karyopherins from this trap [1] can the affinity gradient improve the throughput through the NPC [2, 6].

Direct tests for our model's predictions can now be made by experimentally varying the effective potential experienced by the karyopherin-cargo complexes inside the NPC, by systematically introducing mutations into the binding sites [9], changing the cargo size [10], or using cells with genetically modified numbers of the FG-nups [17]. Finally, any device built according to the principles outlined above would possess the transport properties described by our model, suggesting strategies for creation of artificial nano-molecular sieves.

Materials and methods

The simulations were written in C language and run on a cluster of UNIX processors. The simulation algorithm is described in the text. Analytical calculations were partially performed using Mathematica 5.1 package.

Acknowledgements The authors are thankful to J. Aitchison, S. Bohn, T. Chou, R. Peters, B. Timney, J. Novatt and G. Stolovitzky for helpful comments. This work was supported by NIH grants RR00862 (B.T.C.), GM062427 (M.P.R.), GM071329 (M.P.R. B.T.C, A.G.Z, M.O.M.) and RR022220 (M.P.R., B.T.C.)

References

[1] Macara IG, Transport in and out of the nucleus, (2001), Microbiol. Mol. Biol. Revs. 65, 570-594.

[2] Rout MP, Aitchison JD, Magnasco MO , Chait BT, Virtual gating and nuclear transport: the hole picture (2003), Trends in Cell Biology, 13, 622-628.

[3] Fahrenkrog B, Koser J, Aebi U, The nuclear pore complex: a jack of all trades? (2004), Trends Biochem. Sci., 29, 175-182.

[4] Stewart N, Baker RP, Bayliss R, Clayton L, Grant RP, Littlewood T, Matsuura Y, Molecular mechanism of translocation through nuclear pore complexes during nuclear protein import (2001), FEBS Lett. 498, 145-149.

[5] Suntharalingam M, Wente SR, Peering through the nuclear pore: nuclear pore complex structure, assembly and function., Developmental Cell, (2003), 4, 775-789.

[6] Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao YM, Chait BT, The yeast nuclear pore complex composition, architecture and transport mechanism. (2000) J. Cell. Biol., 148, 635-651.

[7] Kubitscheck U, Grunwald D, Hoekstra A, Rohleder D, Kues T, Siebrasse JP, Peters R. Nuclear transport of single molecules: dwell times at the nuclear pore complex (2005) J. Cell Biol., 168, 233.

[8] Yang W, Gelles J, Musser SM, Imaging of single-molecule translocation through nuclear pore complexes (2004), Proc. Natl. Acad. Sci. 101, 12887-12892.

[9] Bayliss R, Littlewood T, Strawn LA, Stewart M, GLFG and FxFG nucleoporins bind to overlapping sites on importin-beta, J. Biol. Chem. 52, 50597-50606.

[10] Feldherr CM, Akin D, The location of the transport gate in the nuclear pore complex (1997) J. Cell Sci., 110, 3065-3070.

[11] Siebrasse JP, Peters R., Rapid translocation of NTF2 through the nuclear pore of isolated nuclei and nuclear envelopes, EMBO Reports, 3, 887-892.

[12] Pyhtila B, Rexach M, A gradient of affinity for thr karyopherin Kap95p along the yeast nuclear pore complex (2003), J. Biol. Chem., 278, 42699-42709.

[13] Isgro TA, Schulten K, Binding dynamics of isolated nuceoporin repeat regions to importin-beta (2005), Structure, 13, 1869-1879.

[14] Peters R, Translocation through the nuclear pore complex: selecticity and speed by reduction-of-dimensionality (2005), Traffic, 6, 421-427.

[15] Liu SM, Stewart M, Structural basis for the high affinity binding of nucleoporin Nup1 to the Saccharomyces cereviciae importin-beta homologue, Kap95p (2005) J. Mol. Biol., 349, 515-525.

[16] Berg HC, *Random walks in Biology*(1993), Princeton University Press.

[17] Strawn LA, Shen T, Shulga N, Goldfarb D, Wente SR, Minimal nuclear pore complexes define FG repeat domains essential for transport (2004), Nat. Cell Biol. 6, 197-206.

[18] Nachury MV, Weis K, The direction of transport through the nuclear pore can be inverted (1999) Proc. Natl. Acad. Sci., 96, 9622-9627.

[19] Paulillo SM, Phillips E, Koser J, Sauder U, Ullmann K, Powers M, Fahrenkrog B, Nucleoporin domain topology is linked to the transpport status of the nuclear pore complex (2005) J. Mol. Biol, 351, 784-798.

[20] Denning DP, Patel SS, Uversky V, Fink AL, Rexach M, Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded (2003), Proc. Natl. Acad. Sci., 100, 2450-2455.

[21] Lim RYH, Huang NP, Koser J, Deng J, Lau KHA, Schwarz-Herion K, Fahernkrog B, Aebi U, Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport, Proc. Natl. Acad. Sci., 103, 9512-9517.

[22] Gardiner M, *Stochastic Processes in Physics and Chemistry*(2003), Springer-Verlag.

[23] Berezhkovskii AM, Pustovoit MA, Bezrukov SM, Channel-facilitated membrane transport: transit probability and interaction with the channel (2002), J. Chem. Phys., 116, 9952-9956.

[24] Lu D, Grayson P, Schulten K, Glycerol conductance and physical asymmtery of the Escherichia voli glycerol facilitator GlpF (2003), Biophys. J, 85, 2977-2987.

[25] Bortz AB, Kalos MH, Lebowitz JL, New algorithm for for monte-carlo simulation of ising spin systems (1975), J. Comp. Phys., 17, 10-18.

[26] Gillespie DT, General method for numerically simulating stochastic time evolution of couple dchemical reactions (1976), J. Comp. Phys. 22, 403-434.

[27] Le Doussal P, Monthus C, Fisher DS, Random walkers in one-dimensional random environments: Exact renormalization group analysis (1999), Phys. Rev. E, 59, 4795-4840.

[28] Hopfield JJ, Kinetic proofreading-new mechanism for reducing errors in biosynthetic processes requiring high specificity (1974), Proc. Natl. Acad. Sci. 71, 4135-4139.

[29] Ribbeck K, Gorlich D, The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion, (2002), EMBO J., 21, 2664-2771.

[30] Bickel T, Bruinsma R, The nuclear pore complex mystery and anomalous diffusion in reversible gels (2002), Biophysical J. 83, 3079-3087.

[31] Kustanovich T, Rabin Y, Metastable network model of protein transport through nuclear pores (2004), Biophys. J. 86, 2008-2016.

[32] Berezhkovskii AM, Bezrukov SM, Optimizing transport of metabolites through large channels: molecular sieves with and without binding (2005), Biophys. J., 88, L17-L19.

[33] Ben-Efraim I, Gerace L, Gradient of increasing affinity of importin beta for nucleoporins along the pathway of nuclear import (2001), J. Cell Biol., 152, 411-417.

[34] Julicher F, Ajdari A, Prost J, Modeling molecular motors (1997), Revs. Mod. Phys. 69, 1269-1281.

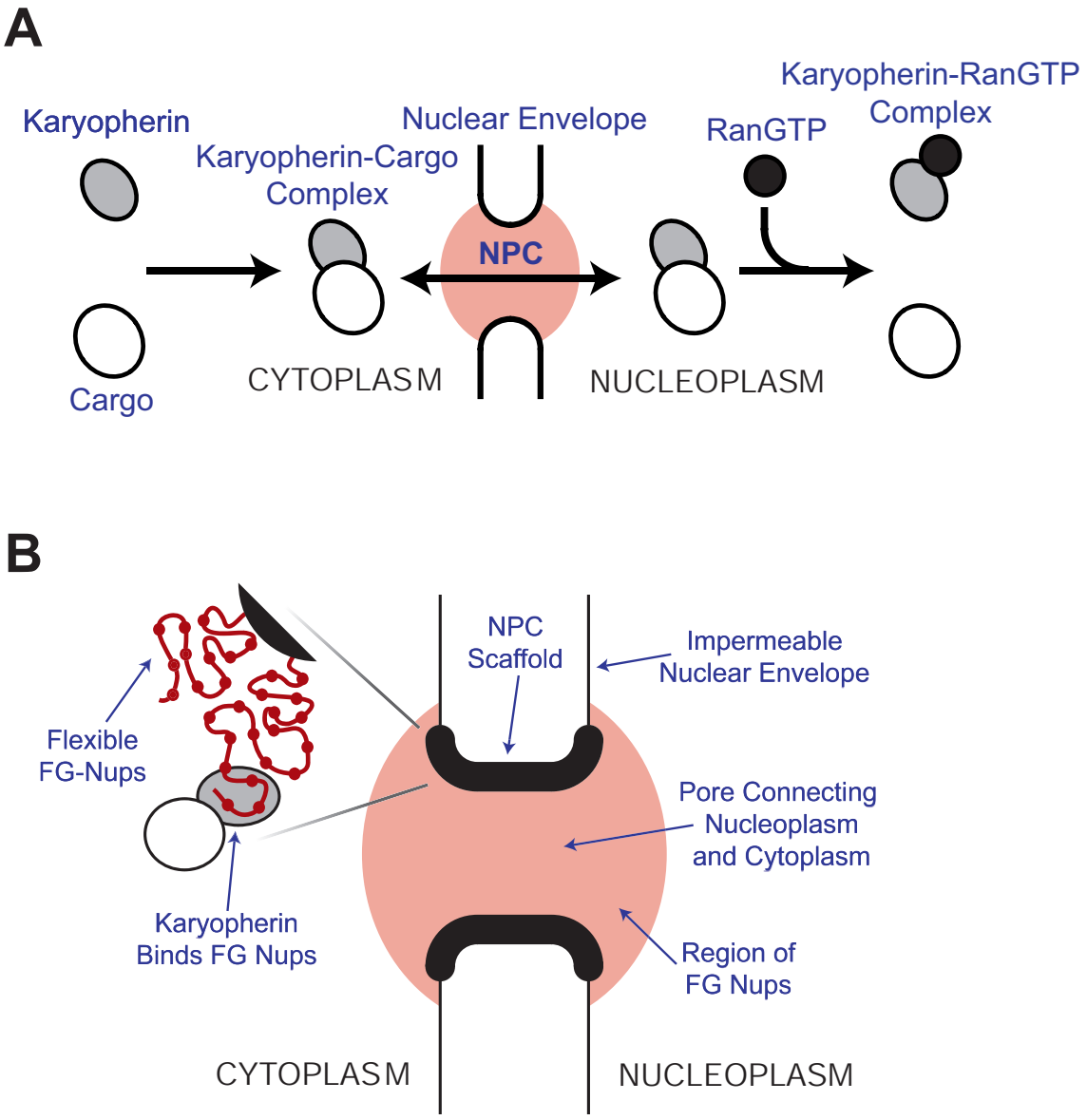

Figure 1: **Main features of NPC and of the nuclear import. A**: Schematic illustration of the nuclear import process. The karyopherins bind the cargo in the cytoplasm and transport it to the nucleus, where the cargo is released by RanGTP. **B**: Diagram of the NPC. The nucleus and the cytoplasm are connected by a channel, which is filled with flexible filaments, FG-nups. The karyopherins carrying a cargo enter from the cytoplasm and hop between the binding sites on the fluctuating FG-nups until they either reach the nuclear side of the NPC, or return to the cytoplasm.

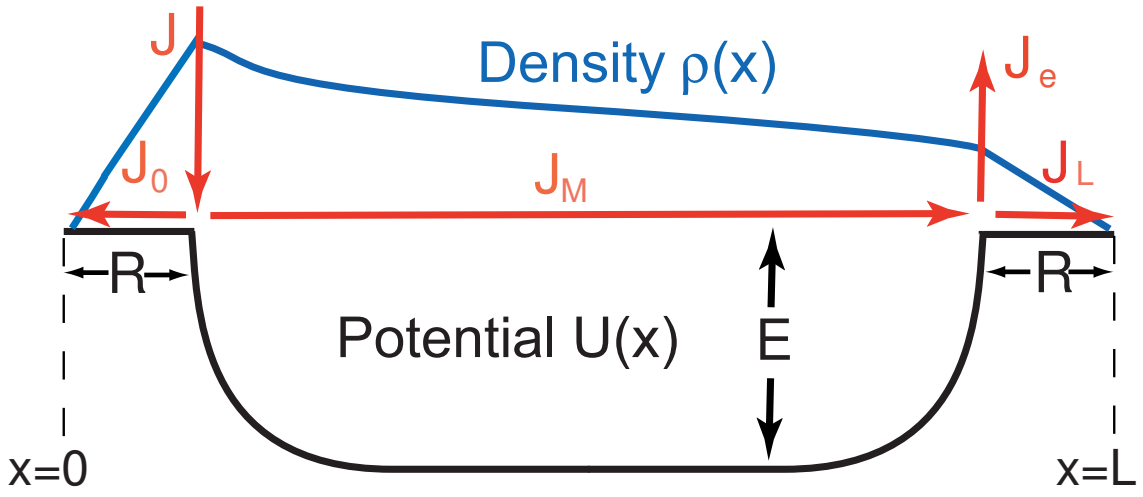

Figure 2: **Transport through the NPC is modeled as diffusion in an energy landscape.** The NPC channel is represented by a potential well $U(x)$, shown in black. The complexes enter the NPC at $x = R$ at an average rate J. A fraction of the entrance flux, J_M, goes through to the nucleus. The rest return to the cytoplasm at an average rate J_0. The exit of the complexes from the channel occurs either due thermal activation, with the rate J_L, or by activated release by RanGTP, with the rate J_e. Steady state particle density inside the channel $\rho(x)$ is shown in blue.

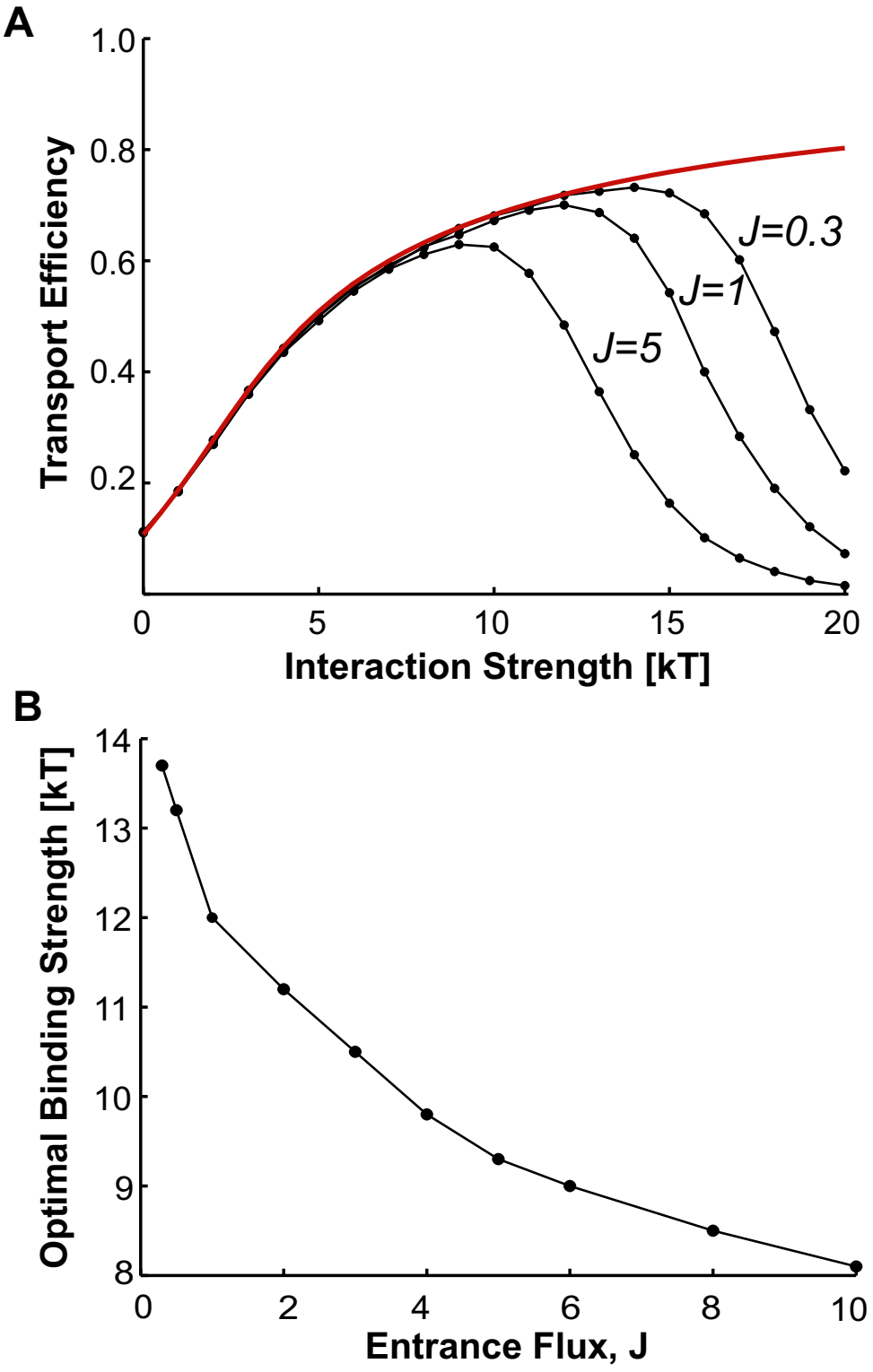

Figure 3: **Transport efficiency is determined by the interaction strength.** **A**: Transport efficiency, as given by probability to reach the nucleus is shown as a function of the interaction strength E. (RanGTP activity in the nucleus is represented by $J_{\rm ran}L^2/(N^2D) = 1.5$). The curves correspond to three different values of the entrance rate J (measured in units of $10^{-4}R^2/(16D)$); the red line is the low rate limit of eq.(4). For any entrance rate, the transport efficiency is maximal at a specific value of the interaction strength, which provides a mechanism of selectivity. **B**: Optimal interaction strength of panel **A** as a function of the incoming rate J (in units of $10^{-4}R^2/(16D)$), for $J_{\rm ran}L^2/(N^2D) = 1.5$. Black dots are simulation results.

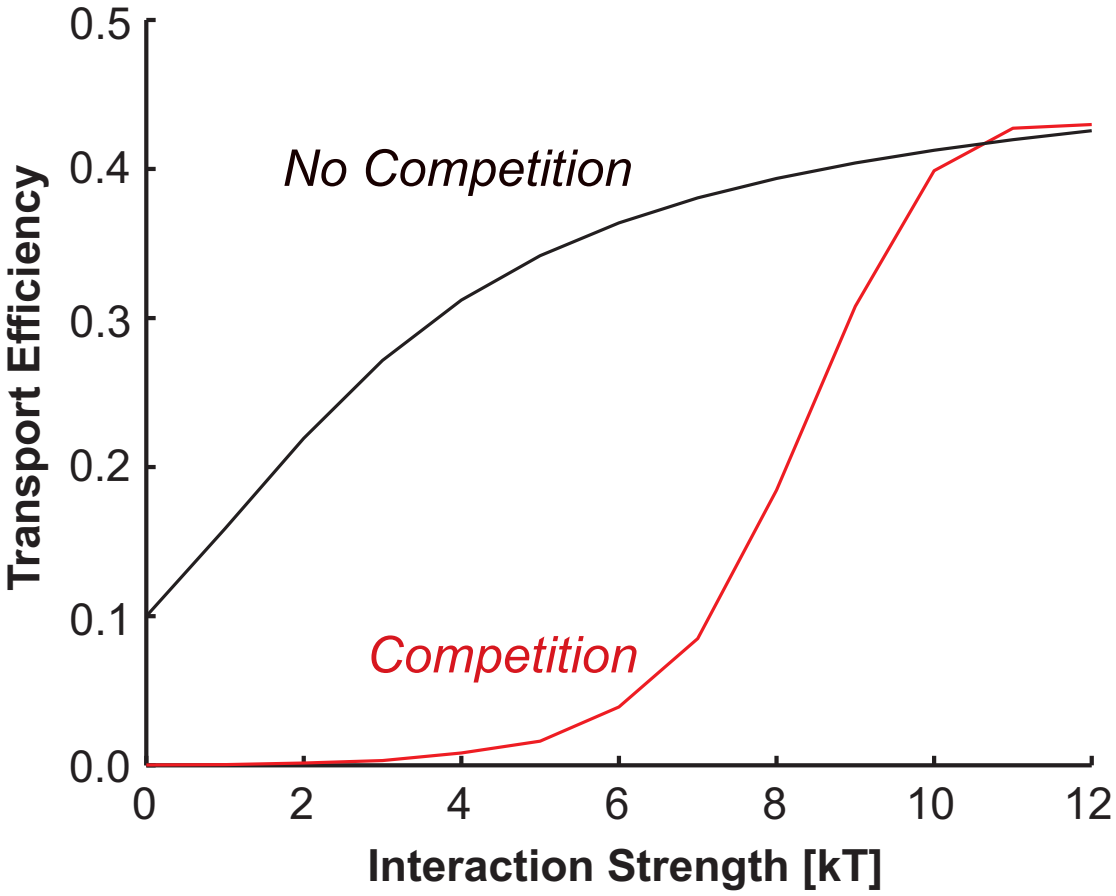

Figure 4: **Karyopherins efficiently exclude non-specifically binding macromolecules from the NPC.** Black line: translocation probability of the complexes as a function of the interaction strength for a single species. Red line: translocation probability of the weakly-binding species in an equal mixture of weakly- and strongly- binding species, as a function of the interaction strength of the weakly-binding species; the interaction strength for the strongly binding species is $12k_BT$. Translocation of the weakly binding species is sharply reduced in the presence of the strongly binding species.

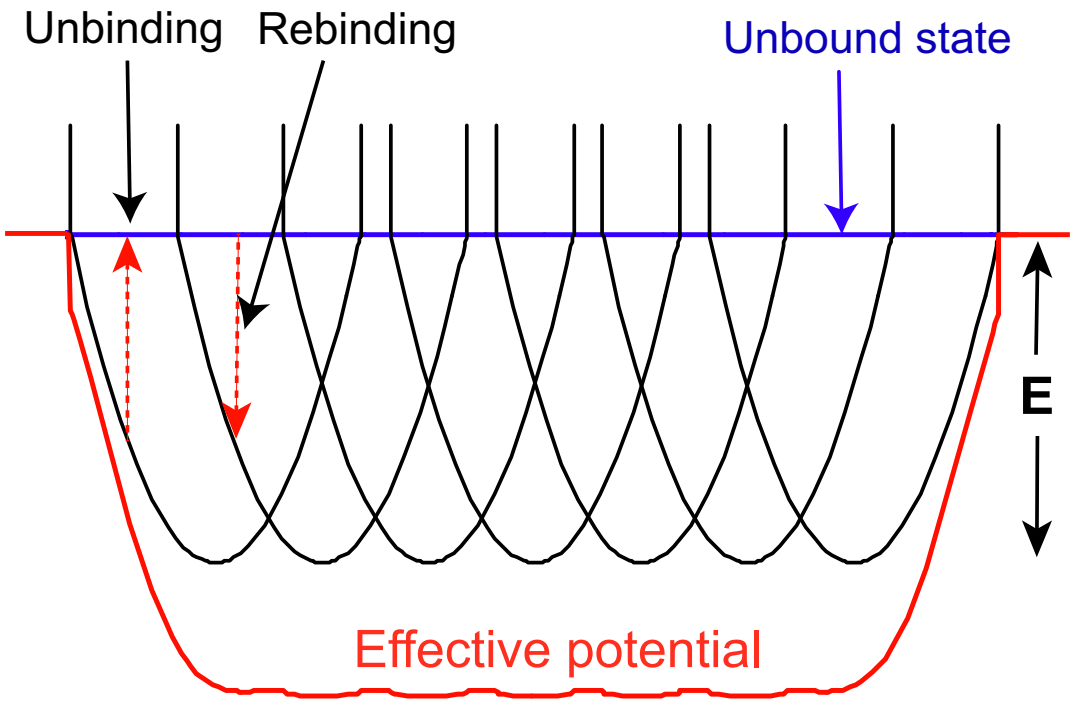

Figure 5: **Discrete overlapping FG-nups can be approximated by a smooth effective potential.** Transport through the NPC can be represented as diffusion in an array of potential wells that represent flexible FG-nups whose fluctuation regions overlap. The red dotted lines correspond to the complexes unbinding from and rebinding to the FG-nups. Blue line is the unbound state. Red line shows the equivalent potential in the case when the unbinding of the complexes from the FG-nups is much faster than the lateral diffusion across an individual well.

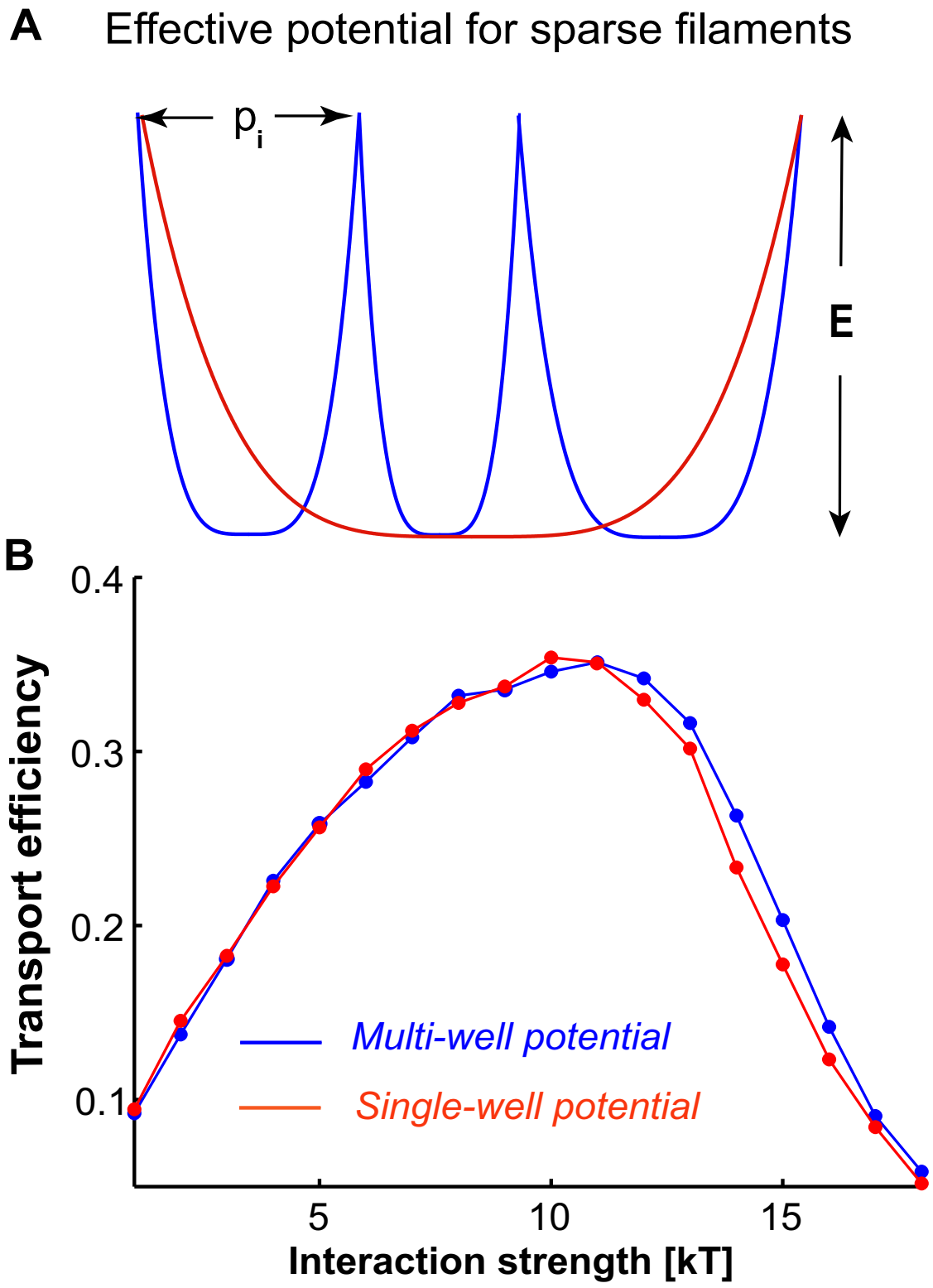

Figure 6: **The number of the FG-nups does not significantly affect the transport properties.** **A**: Effective potential for sparse flexible FG-nups is shown in blue line. Each well corresponds to an FG-nup. The transport properties in this multi-well potential are *independent* of the number of wells, and hence equivalent to those in the single-well potential, shown in red line. **B**: Numerical simulations show essentially identical transport efficiencies in a multi-well potential (blue line) and in the single-well potential (red line); $U_0(x) = ((x/p)^2 - 1)$

Supporting information

Residence time in a multi-well potential

Here we calculate the residence time in the potential of Fig. 5. As before, we assume that the potential is zero in the intervals $[0, R]$ and $[L-R, L]$. The particles start their diffusion from $x = R$. For an arbitrary distribution of wells widths p_i, the potential is equal to $U_0(x/p_i)$ in an interval $x_i < x < x_{i+1}$, where $x_i = \sum_{j<i} p_j$. The mean first passage time, until a particle either reaches to $x = L$ or returns to $x = 0$ (with $J_e = 0$, is:

$$T(U)/D = P_{tr}(U)T_{\rightarrow}(U) + (1 - P_{tr}(U))T_{\leftarrow}(U) \tag{9}$$

where $T_{\rightarrow}(U) = \int_R^L dy e^{U(y)} \int_R^y dz e^{-U(z)}$; $T_{\leftarrow}(U) = \int_0^R dy e^{U(y)} \int_y^R dz e^{-U(z)}$ and $P_{tr} = \int_0^R e^{U(x)} dx / \int_0^L e^{U(x)}$ is the translocation probability [22].

Thus, $T_{\leftarrow} = R^2/2$ and

$$T_{\rightarrow}(U) = \frac{1}{2}R^2 + R\int_0^{p_0} e^{-U(x)}dx + \int_0^{p_0} e^{U(y)}dy \int_0^y e^{-U(x)}dx$$

The second and the third terms in the above expression become, respectively:

$$\begin{aligned} I_1 &= \sum_{i=1}^{n} \int_{x_i}^{x_{i+1}} e^{U_0(\frac{x}{p_i})}dx = p_0 \int_0^1 e^{U_0(u)}du \\ I_2 &= \sum_{i=1}^{n} \int_0^{p_i} e^{U_0(\frac{y}{p_i})}dy \left[\int_0^y e^{-U_0(\frac{u}{p_i})}du + \sum_{j=1}^{i-1} \int_0^{p_j} e^{-U_0(\frac{u}{p_j})}du\right] \\ &= (\sum_{i=1}^{n} p_i^2 + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} p_i p_j) I_0 = p_0^2 I_0 \end{aligned}$$

where $I_0 = \int_0^1 e^{U_0(x)}dx \int_0^x e^{-U_0(y)}dy = \frac{1}{2}\int_0^1 \int_0^1 dxdy e^{U_0(x)}e^{-U_0(y)}$. Thus, the mean residence time $T(U)$ does not depend on the number of wells n.

The effect of the additional exit current J_e at $x = L-R$ is formally equivalent to shortening the interval $[L-R, L]$ to $[L - R/(1+K), L]$. Repeating the calculation above for this case shows that also for non-zero J_e the residence time does not depend on the number of wells.